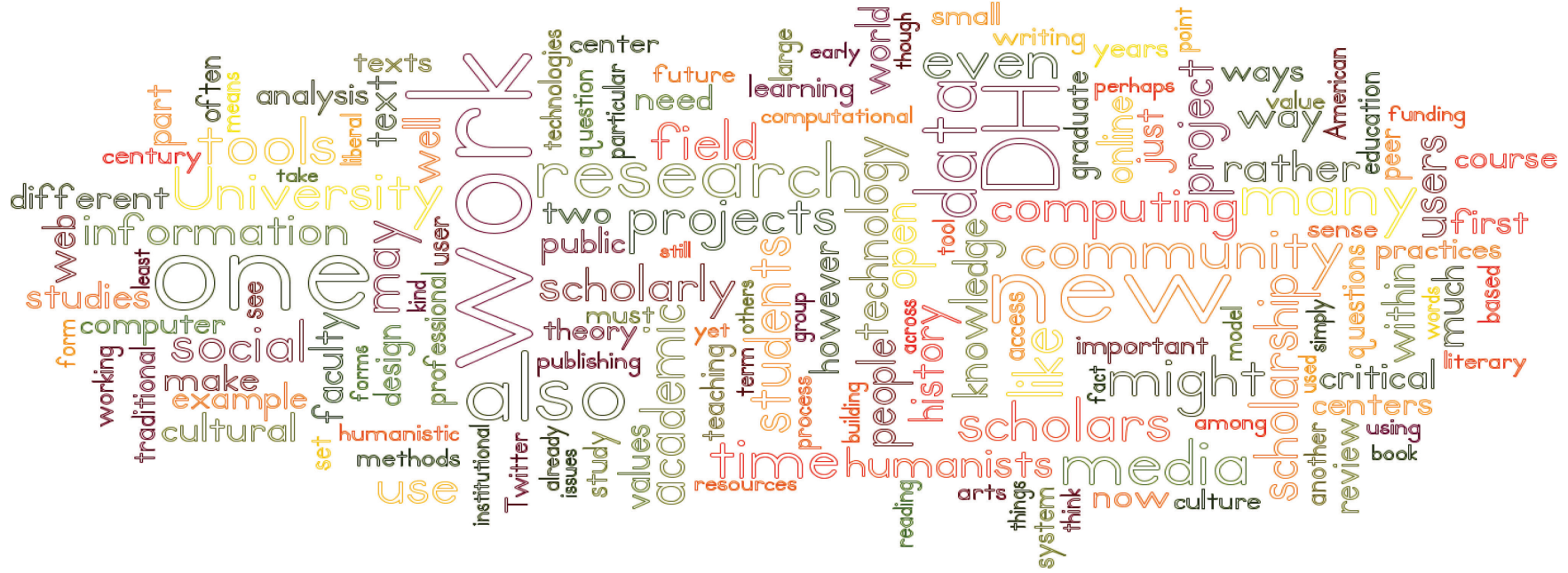


Visualizing Scholarly Communication

Nina Collins and Matthew Hannah
Purdue University, West Lafayette

What is Digital Humanities?



Wordcloud of most frequently used words in *Debates in the Digital Humanities*

What is “Digital Humanities”?

“[T]he field of digital humanities is characterized at a deeper level by a more critical engagement with technology” (Julia Flanders, *Defining DH*, 205).

“I think of digital humanities as an umbrella term that covers a wide variety of digital work in the humanities” (Kathie Gossett, *Debates in DH*, 67).

“Digital humanities is a spectrum” (Lincoln Mullen, *Defining DH*, 237).

“Personally, I think Digital Humanities is about building things” (Stephen Ramsay, *Defining DH*, 241).

“A term of tactical convenience” (Matthew Kirschenbaum, *Debates in DH*, 68).

“The use of digital tools and methods in humanities study and dissemination” (Geoffrey Rockwell, *Debates in DH*, 69).



Image retrieved from: Wikipedia entry for “Computing in the Soviet Union”

Methods for visualizing analyzing scholarly communication...

DH offers tools and methods to address scholarly questions.

Possible methods:

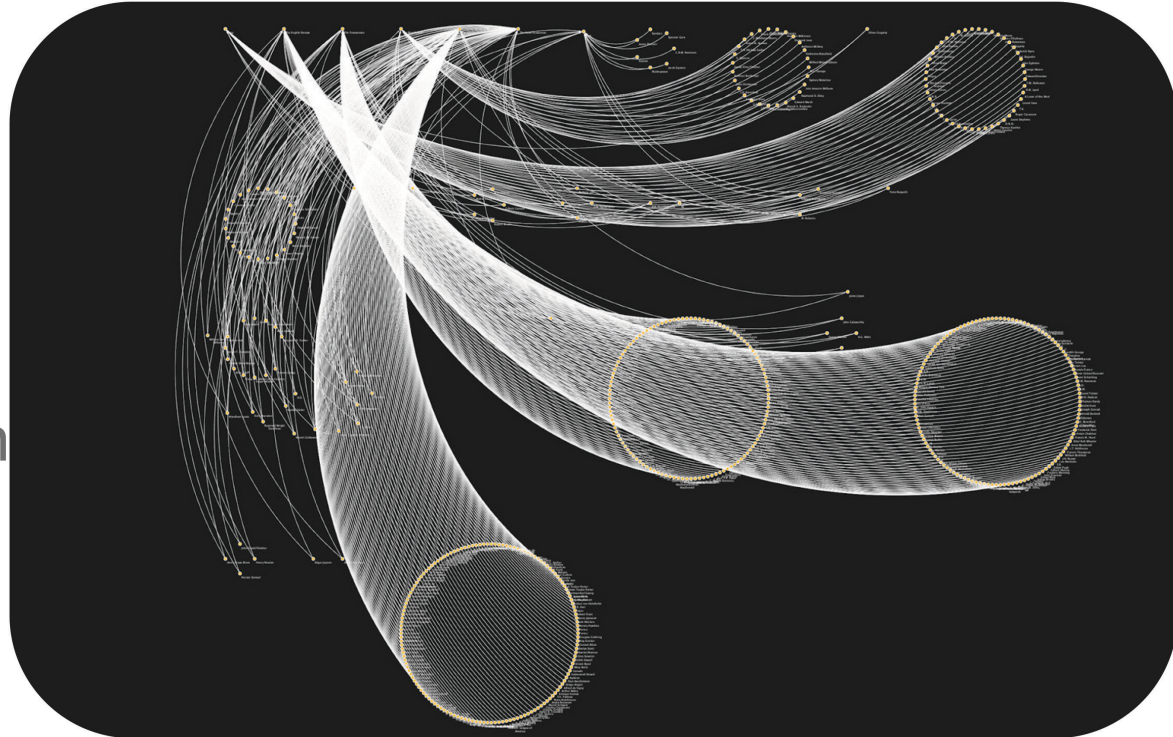
Text analysis

Network analysis

Data Visualization

Bibliometrics

Data Science



DH as collaborator with Scholarly Communications

- Collaboration is key to academic work.
- Offers range of digital tools to analyze or visualize results. Some tools may be familiar, but some may be new.
- Bridges to other areas of research, including bibliometrics but also to text analysis, data mining, data science, and other methods that could be reimagined.
- Critical apparatus to translate research to humanities/social sciences context but also apply humanistic forms of critique or analysis to scholarly communications.

BEALL'S LIST *of*
PREDATORY
open
access
PUBLISHERS



Notes for previous slide:

Mr. Jeffrey Beall is responsible for starting the discussion of predatory journals and publishers. He wanted to draw attention to publishing practices that he felt were sub-par. He identified journals that did not adhere to publishing best practice--such as not performing adequate peer review when they claimed to do so. He called them predatory. Those who did not use transparent business practice were labelled predatory (only if they were also open access). Journals who engaged in deceptive marketing and business practices were also labelled predatory--again, if they were open access journals.

He started a now-famous blog, where he identified specific open access journals or publishers as “predatory”. It is interesting to note that users of his site did not have access to his scoring criteria for any given journal on his list. Many critics believe a journal would make Mr. Beall’s “predatory” list simply because the journal was OA, or simply because Mr. Beall decreed it predatory. While he was clear to point out the lack of transparency of the journals of which he was so critical, he did not practice a fair degree of transparency on his own site. This site became the authoritative go-to resource to determine whether or not an unfamiliar open access journal or publisher should be trusted. It is interesting to note that much of Mr. Beall’s publications, decrying the fallacy of these journals who do not actually peer review the content within them, was published in opinion or commentary pieces--themselves not peer reviewed. I would like to make clear that I am not bringing this up to paint Mr. Beall in a bad light, so much as I would like to point out that Beall’s list of Predatory Open Access Publishers is still perceived by many to be an authority on this subject, meant to bring awareness and call for using a critical lense to evaluate journals and publishers. I find it ironic, at best, that the same critique was not applied (by the scholarly community) to the blog “publication” of Mr. Beall.

Predatory Publishing

Umbrella term:

Scientific misconduct, lack of peer review, deception, business fraud, citation manipulation,

Current Scholarship doesn't clearly define what constitutes predatory publishing

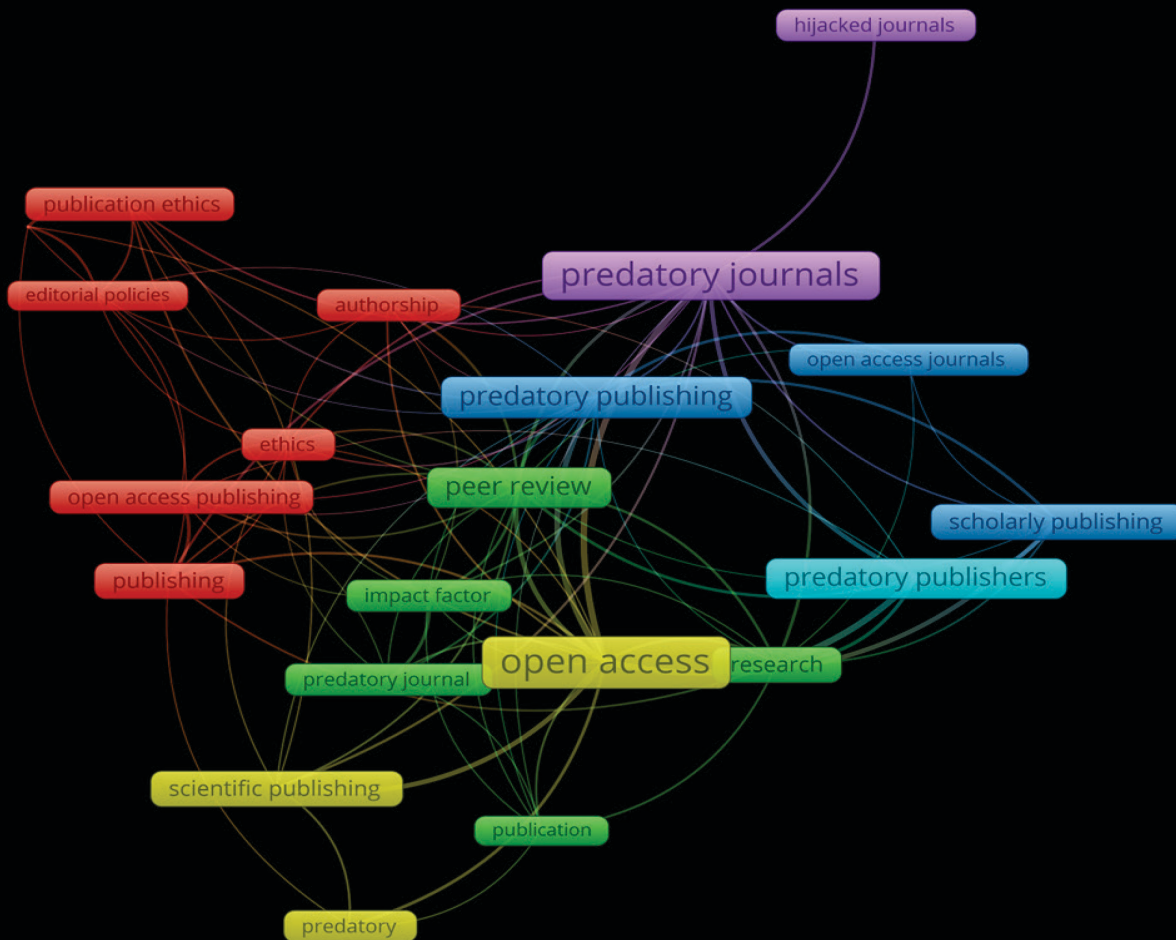
Unhelpful term blinds us to challenges and opportunities in publishing. Journals are not well served by being labelled “predatory” or “not predatory”.

Our method... VosViewer

Created a list pulled from Scopus of keywords from all articles featuring the keywords “predatory journal*” or “predatory publishing*”

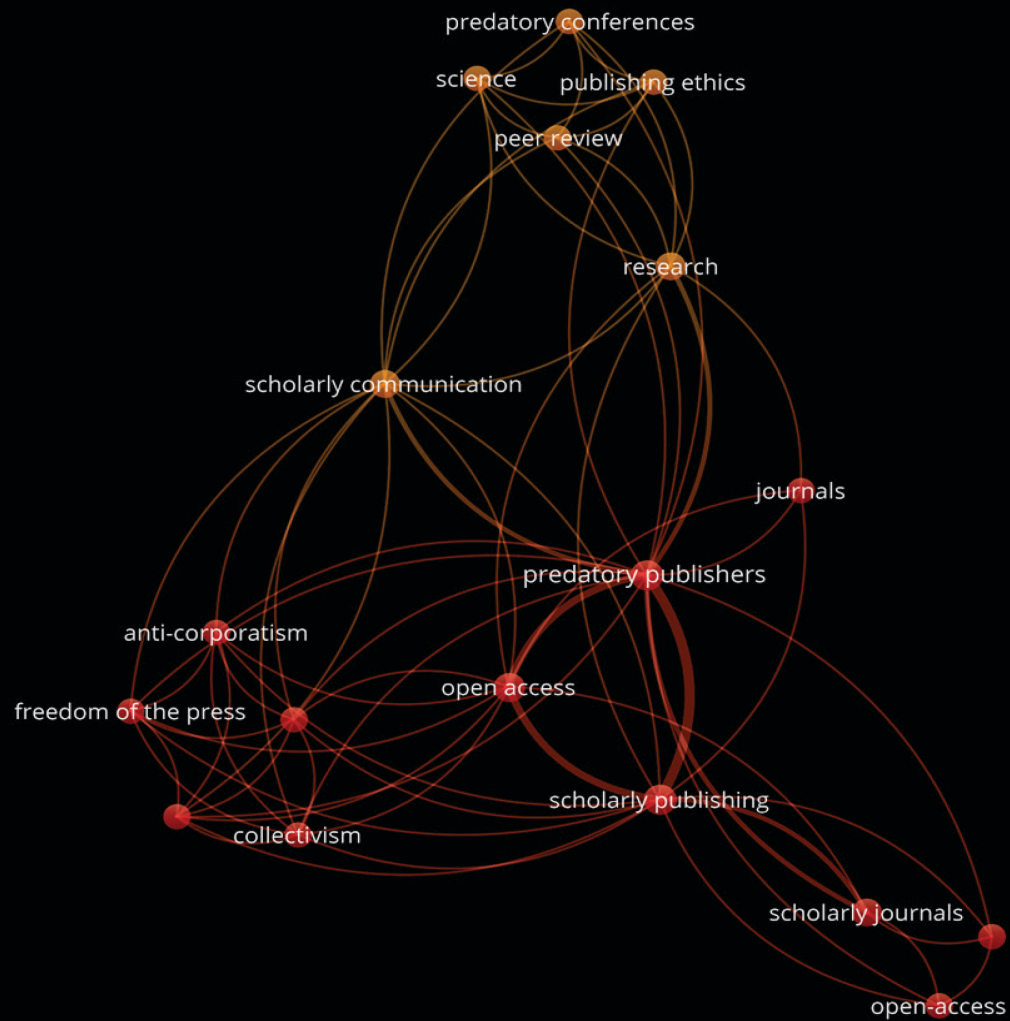
Used VosViewer to create a network showing links between keywords. Colors and grouping represent communities where mutual linkages are more common.

Networks visualize large systems of nodes connected by edges to show the larger bibliographic relationships.



Notes for previous slide:

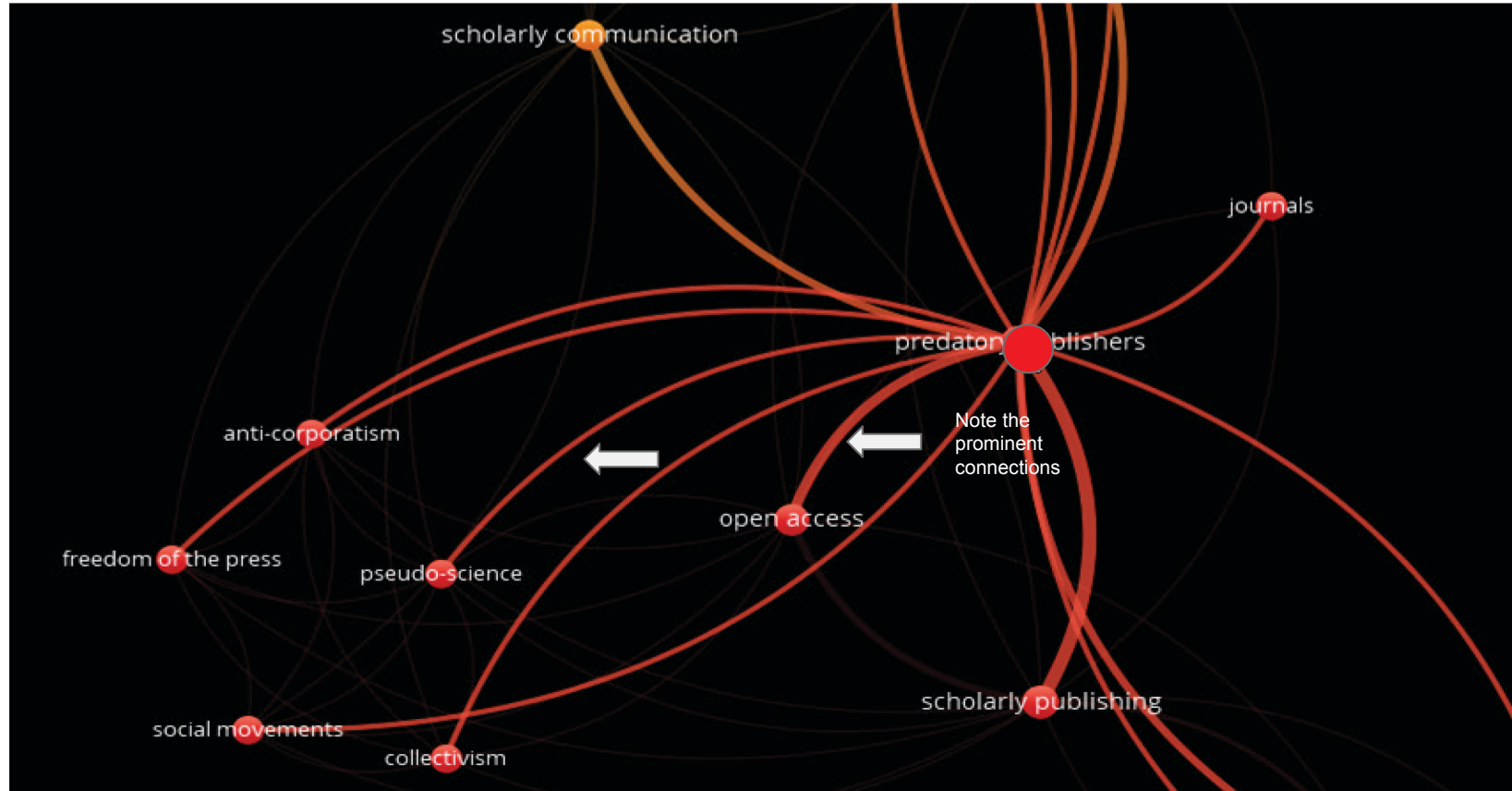
This image represents the interconnections of the different keywords that appear in the bibliographic export. Note the interconnected topics.



Notes for previous slide:

Next, we drilled down on Jeffrey Beall's publication record to compare data. Jeffrey Beal's published articles, focusing on keyword co-occurrence related to keywords "predatory" "journal*" and "publishing*"

Close up on keywords from all articles related to Jeffrey Beall's work on predatory publishing.



Our method...Voyant

- Voyant Tools allows scholars to conduct text analysis on bodies of text, such as abstracts or full articles.
- First, we scraped abstracts from articles about predatory publishing and predatory journals, and we visualized the results. We pasted the abstracts into Voyant Tools to generate a visualization of the most common concepts in such texts.
- Then, we gathered the central corpus of Beall's writing on predatory publishing in order to visualize the larger themes in his work.
- Comparing these two collections will show what the major concepts in the field are and allow us to compare them to Beall's written corpus.

Visualizing “predatory open access”...



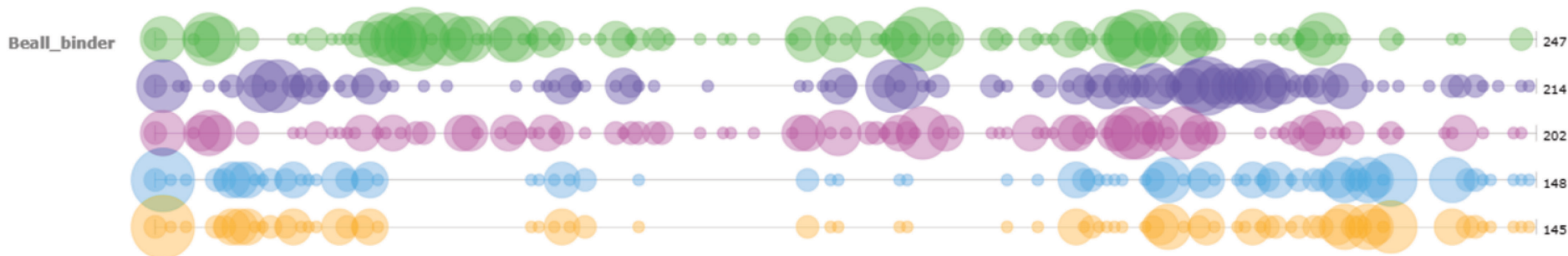
Most frequently used words in Beall's works.



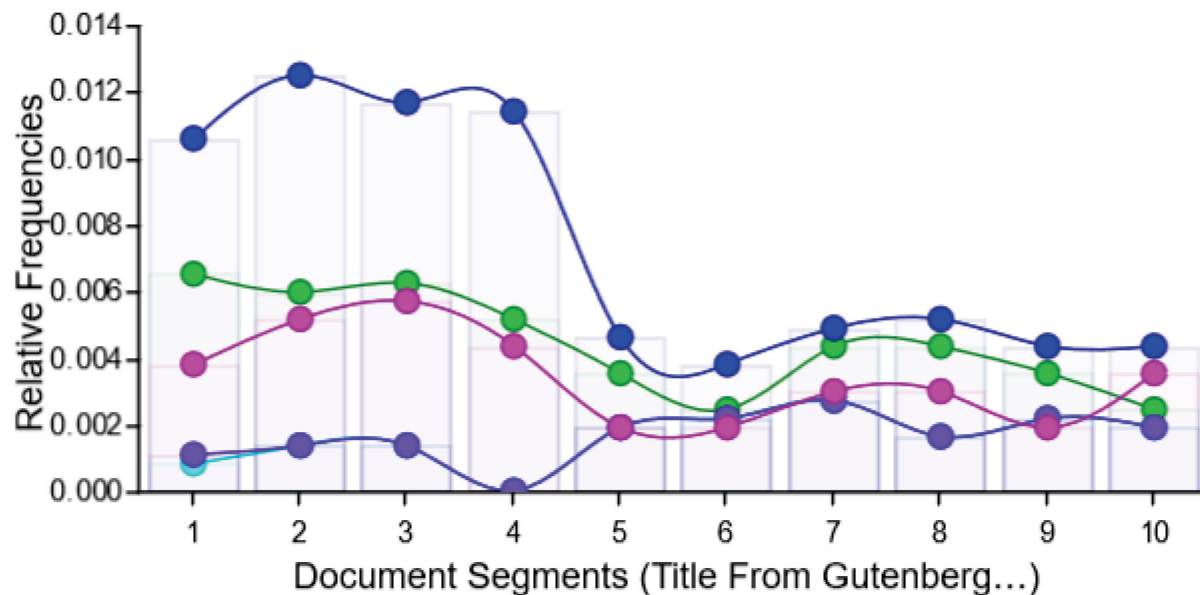
Most frequently used words in Beall's titles and titles of works that have cited Beall's works.

Topical Trends in Beall

Open, Publishers, Access, Predat*, Predatory



Our results...Voyant



Visualization showing the most common terms listed 364 titles related to predatory publishing.

Conclusions and future directions...

Digital Humanities tools and techniques can be applied in so many situations.

Our findings:

- OA advocates: we have our work ahead of us to disambiguate “open access” from “predatory publishers”.
- The term “predatory” is unhelpful to publishing as well as OA.
- As this unhelpful term is getting wider attention, we need to be proactive. No longer ignore the issue and assume it's not our problem.

Future directions:

- While we looked at the literature on “predatory journals” to see the relatedness of open access, we did not look at the much larger corpus of literature on open access to see keyword associations.